Uncertainty Quantification and Chemical Model Reduction

Habib N. Najm

hnnajm@sandia.gov

Combustion Research Facility
Sandia National Laboratories, Livermore, CA

Multi Agency Coordination Committee for Combustion Research
SNL-CRF
Sep. 17-20, 2012

Acknowledgement

B.J. Debusschere, R.D. Berry, K. Sargsyan, C. Safta,
K. Chowdhary — Sandia National Laboratories, CA

R.G. Ghanem — U. South. California, Los Angeles, CA
O.M. Knio   — Duke Univ., Durham, NC
O.P. Le Maître — CNRS, Paris, France
Y.M. Marzouk — Mass. Inst. of Tech., Cambridge, MA

Outline

1. **Motivation**

2. Uncertainty Quantification – Basics
   - Parameter Estimation
   - Forward Propagation of Uncertainty

3. Analysis and Reduction of Uncertain Chemical Models

4. Closure

The Case for Uncertainty Quantification (UQ)

UQ is needed in:

- Assessment of confidence in computational predictions
- Validation and comparison of scientific/engineering models
- Design optimization
- Use of computational predictions for decision-support
- Assimilation of observational data and model construction
- Multiscale and multiphysics model coupling

UQ Relevance in Chemical Model Reduction

- Detailed chemical models have significant uncertainty
    - Model structure
        - – Species and reactions
        - – Reaction rate expressions
        - – Enthalpies/entropies of formation $f(T)$
    - Model parameters
        - – Reaction rate coefficients
        - – Enthalpies/entropies of formation coefficients
- Reduced chemical models need to perform over the range of given uncertainties – robustness
- Need to consider a comprehensive error budget:
  $$\{\|M_{\text{reduced}} - M_{\text{detailed}}\|, \text{ uncertainty}\}$$
    - may justify more lenient error requirements on reduced models

Overview of UQ Methods

Estimation of model/parametric uncertainty

- Expert opinion, data collection
- Regression analysis, fitting, parameter estimation
- Bayesian inference of uncertain models/parameters

Forward propagation of uncertainty in models

- Local sensitivity analysis (SA) and error propagation
- Fuzzy logic; Evidence theory — interval math
- Probabilistic framework — Global SA / stochastic UQ
    - Random sampling, statistical methods
    - Galerkin methods
        - Polynomial Chaos (PC) — intrusive/non-intrusive
    - Collocation methods: PC/other interpolants — non-intrusive

Parameter Estimation

- Model calibration — Inverse problem – Bayes rule
  - Data:
    - Experimental observations
    - Computational predictions – high-fidelity "truth" model
  - Missing data — Bayesian Imputation
    - – Simulates missing data using posterior predictive distribution given observed values
    - – Observed data posterior
  - No data – but given summary statistics
    - Simulate data satisfying summary statistics/constraints
    - Pooled posterior
- Expert elicitation
- Computational predictions — Forward UQ

Bayes formula for Parameter Inference

- Data Model (fit model + noise model): $\quad y = f(\lambda) * g(\epsilon)$
- Bayes Formula:

$$p(\lambda, y) \;=\; p(\lambda|y)p(y) = p(y|\lambda)p(\lambda)$$

$$\underset{\text{Posterior}}{p(\lambda|y)} = \frac{\overset{\text{Likelihood}}{p(y|\lambda)}\;\overset{\text{Prior}}{p(\lambda)}}{\underset{\text{Evidence}}{p(y)}}$$

- Prior: knowledge of λ prior to data
- Likelihood: forward model and measurement noise
- Posterior: combines information from prior and data
- Evidence: normalizing constant for present context

Prior Modeling

- Informative prior
- (Mostly) Uninformative prior
  - Improper prior
  - Objective prior
  - Maxent prior
  - Reference prior
  - Jeffreys prior
- The choice of prior can be crucial when there is little information in the data relative to the number of degrees of freedom in the inference problem
- When there is sufficient information in the data, the data can overrule the prior

Likelihood Modeling

- This is frequently the *core* modeling challenge
  - Error model: a statistical model for the discrepancy between the forward model and the data
  - composition of the error model with the forward model
- Hierarchical Bayes modeling, and dependence trees

$$p(\phi, \theta) = p(\phi|\theta)p(\theta)$$

- Choice of observable – constraint on Quantity of Interest?
- Stochastic versus Deterministic forward models
  - Intrinsic noise term, *e.g.* Langevin eqn.
  - Specified uncertain parameter in fit model
- Error model composed of discrepancy between
  - data and the truth – (data error)
  - model prediction and the truth – (model error)

  - Mean bias and correlated/uncorrelated noise structure

Experimental Data

- Empirical data error model structure can be informed based on knowledge of the experimental apparatus
- Both bias and noise models are typically available from instrument calibration
- Noise PDF structure
    - A counting instrument would exhibit Poisson noise
    - A measurement combining many noise sources would exhibit Gaussian noise
    - Noise correlation structure
        – Point measurement
        – Field measurement
- Error model composed of model error + data error

Exploring the Posterior

- Given any sample λ, the un-normalized posterior probability can be easily computed

$$p(\lambda|y) \propto p(y|\lambda)p(\lambda)$$

- Explore posterior w/ Markov Chain Monte Carlo (MCMC)
    - Metropolis-Hastings algorithm:
        - Random walk with proposal PDF & rejection rules
    - Computationally intensive, $\mathcal{O}(10^5)$ samples
    - Each sample: evaluation of the forward model
        - Surrogate models
- Evaluate moments/marginals from the MCMC statistics

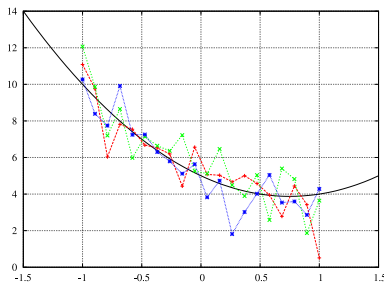# Bayesian inference illustration: noise$\uparrow$ $\Rightarrow$ uncertainty$\uparrow$
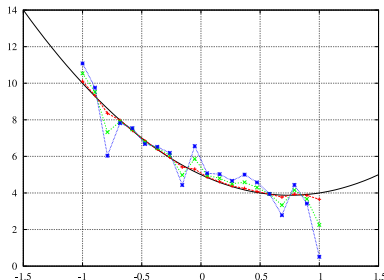
- data: $y = 2x^2 - 3x + 5 + \epsilon$
- $\epsilon \sim \mathcal{N}(0, \sigma^2)$, $\sigma = \{0.1, 0.5, 1.0\}$
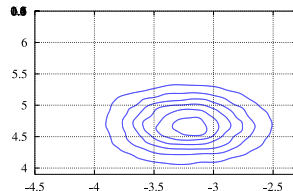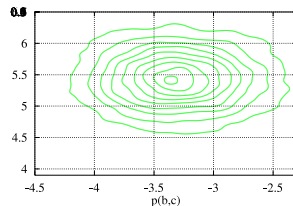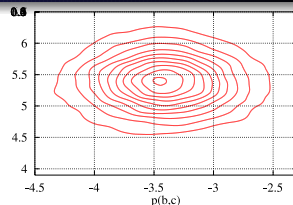- Fit model $y = ax^2 + bx + c$

Marginal posterior density $p(a, c)$:
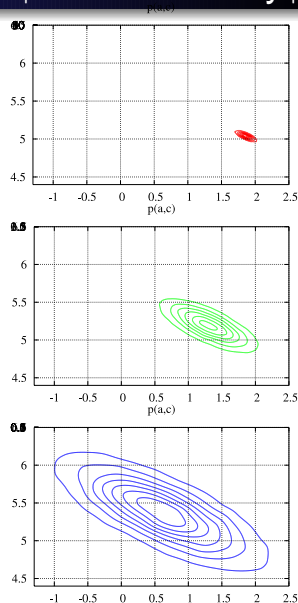
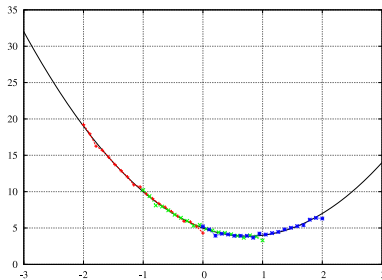# Bayesian illustration: Data realization ⇒ posterior

- data: $y = 2x^2 - 3x + 5 + \epsilon$
- $\epsilon \sim \mathcal{N}(0, 1)$
  - 3 different random seeds
- Fit model $y = ax^2 + bx + c$

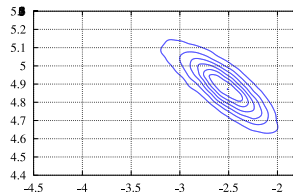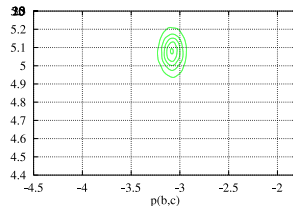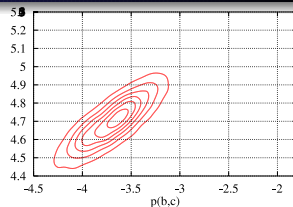Marginal posterior density $p(b, c)$:

Illustration: Data range $\Rightarrow$ correlation structure

- data: $y = 2x^2 - 3x + 5 + \epsilon$
- $\epsilon \sim \mathcal{N}(0, 0.04)$
- ranges: $x \in \{[-2, 0], [-1, 1], [0, 2]\}$
- Fit model $y = ax^2 + bx + c$

Marginal posterior density $p(b, c)$:

Probabilistic Forward UQ & Polynomial Chaos Representation of Random Variables

- With $y = f(x)$, x a random variable, estimate the RV y
- Can describe a RV in terms of its density, moments, characteristic function, or most fundamentally as a function on a probability space
- Constraining the analysis to RVs with finite variance, enables the representation of a RV as a spectral expansion in terms of orthogonal functions of standard RVs.
    - Polynomial Chaos
- Enables the use of available functional analysis methods for forward UQ

Polynomial Chaos Expansion (PCE)

- Model uncertain quantities as random variables (RVs)
- Given a *germ* $\boldsymbol{\xi}(\omega) = \{\xi_1, \cdots, \xi_n\}$ – a set of *i.i.d.* RVs
  - where $p(\boldsymbol{\xi})$ is uniquely determined by its moments

Any RV in $L^2(\Omega, \mathfrak{S}(\boldsymbol{\xi}), P)$ can be written as a PCE:

$$u(\boldsymbol{x}, t, \omega) = f(\boldsymbol{x}, t, \boldsymbol{\xi}) \simeq \sum_{k=0}^{P} u_k(\boldsymbol{x}, t) \Psi_k(\boldsymbol{\xi}(\omega))$$

- $u_k(\boldsymbol{x}, t)$ are mode strengths
- $\Psi_k()$ are functions orthogonal w.r.t. $p(\boldsymbol{\xi})$

With dimension n and order p: $\qquad P + 1 = \dfrac{(n + p)!}{n! p!}$

Orthogonality

By construction, the functions $\Psi_k()$ are orthogonal with respect to the density of $\boldsymbol{\xi}$

$$u_k(\boldsymbol{x}, t) \;=\; \frac{\langle u\Psi_k \rangle}{\langle \Psi_k^2 \rangle} = \frac{1}{\langle \Psi_k^2 \rangle} \int u(\boldsymbol{x}, t; \lambda(\boldsymbol{\xi}))\, \Psi_k(\boldsymbol{\xi}) p_{\boldsymbol{\xi}}(\boldsymbol{\xi}) d\boldsymbol{\xi}$$

Examples:

- Hermite polynomials with Gaussian basis
- Legendre polynomials with Uniform basis, ...
- Global versus Local PC methods
    - Adaptive domain decomposition of the support of $\boldsymbol{\xi}$

Essential Use of PC in UQ

Strategy:

- Represent model parameters/solution as random variables
- Construct PCEs for uncertain parameters
- Evaluate PCEs for model outputs

Advantages:
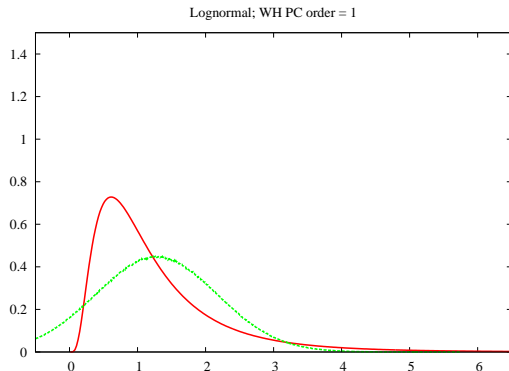
- Computational efficiency
- Sensitivity information

Requirement:

- Random variables in L^2, i.e. with finite variance

PC Illustration: WH PCE for a Lognormal RV

- Wiener-Hermite PCE constructed for a Lognormal RV
- PCE-sampled PDF superposed on true PDF
- Order = 1

$$u = \sum_{k=0}^{P} u_k \Psi_k(\xi)$$
$$= u_0 + u_1 \xi$$

Lognormal; WH PC order = 1

PC Illustration: WH PCE for a Lognormal RV

- Wiener-Hermite PCE constructed for a Lognormal RV
- PCE-sampled PDF superposed on true PDF
- Order = 2

Lognormal; WH PC order = 2

$$u = \sum_{k=0}^{P} u_k \Psi_k(\xi)$$

$$= u_0 + u_1\xi + u_2(\xi^2 - 1)$$

PC Illustration: WH PCE for a Lognormal RV

- Wiener-Hermite PCE constructed for a Lognormal RV
- PCE-sampled PDF superposed on true PDF
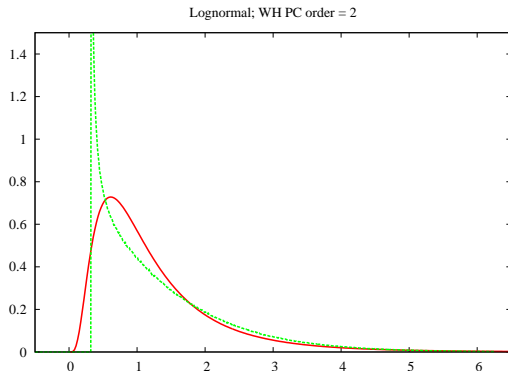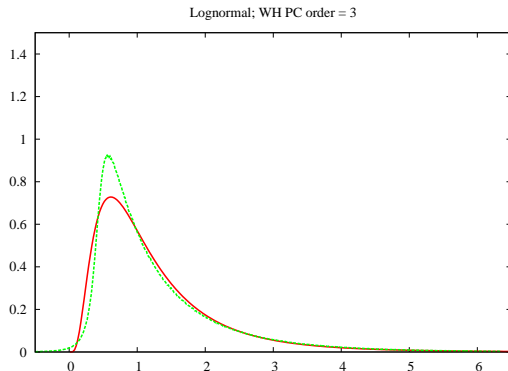- Order = 3

Lognormal; WH PC order = 3

$$u = \sum_{k=0}^{P} u_k \Psi_k(\xi)$$
$$= u_0 + u_1\xi + u_2(\xi^2 - 1) + u_3(\xi^3 - 3\xi)$$

PC Illustration: WH PCE for a Lognormal RV

Lognormal; WH PC order = 4

- Wiener-Hermite PCE constructed for a Lognormal RV
- PCE-sampled PDF superposed on true PDF
- Order = 4

$$u = \sum_{k=0}^{P} u_k \Psi_k(\xi)$$

$$= u_0 + u_1\xi + u_2(\xi^2 - 1) + u_3(\xi^3 - 3\xi) + u_4(\xi^4 - 6\xi^2 + 3)$$
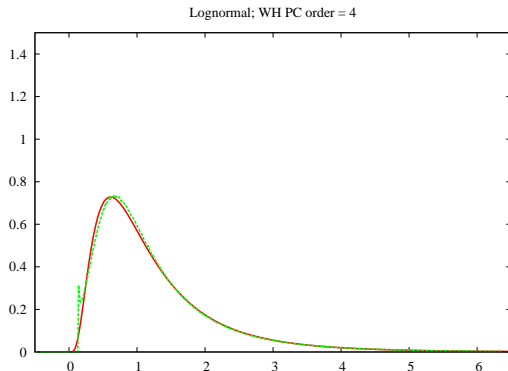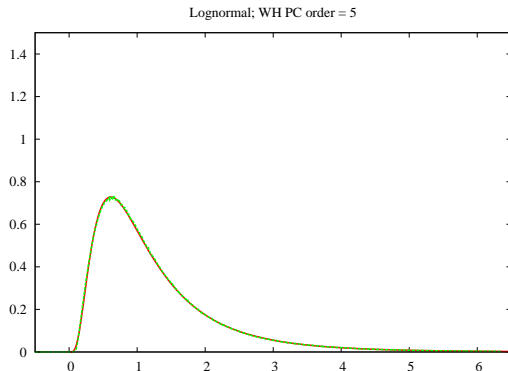
PC Illustration: WH PCE for a Lognormal RV

- Wiener-Hermite PCE constructed for a Lognormal RV
- PCE-sampled PDF superposed on true PDF
- Order = 5

Lognormal; WH PC order = 5
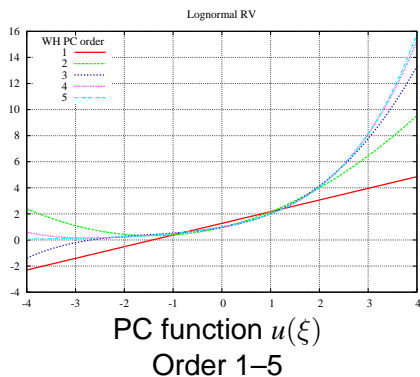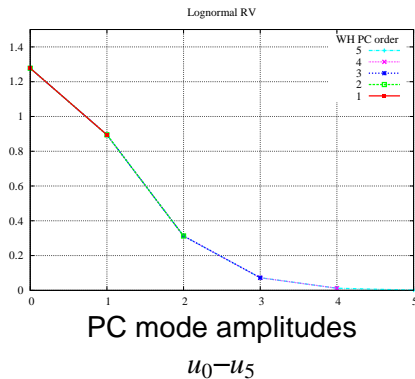
$$u = \sum_{k=0}^{P} u_k \Psi_k(\xi)$$

$$= u_0 + u_1\xi + u_2(\xi^2 - 1) + u_3(\xi^3 - 3\xi) + u_4(\xi^4 - 6\xi^2 + 3) + u_5(\xi^5 - 10\xi^3 + 15\xi)$$

PC Illustration: WH PCE for a Lognormal RV

PC mode amplitudes
$u_0 - u_5$

PC function $u(\xi)$
Order 1–5

- Fifth-order Wiener-Hermite PCE represents the given Lognormal well
- Higher order terms are negligible

Intrusive PC UQ: A direct *non-sampling* method

- Given model equations: $\mathcal{M}(u(\boldsymbol{x}, t); \lambda) = 0$

- Express uncertain parameters/variables using PCEs

$$u = \sum_{k=0}^{P} u_k \Psi_k; \quad \lambda = \sum_{k=0}^{P} \lambda_k \Psi_k$$
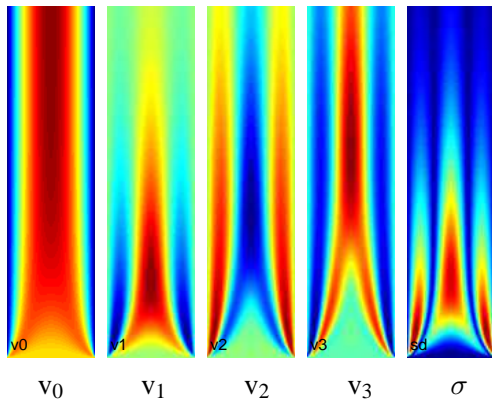
- Substitute in model equations; apply Galerkin projection

- New set of equations: $\mathcal{G}(U(\boldsymbol{x}, t), \Lambda) = 0$
  - with $U = [u_0, \ldots, u_P]^T$, $\Lambda = [\lambda_0, \ldots, \lambda_P]^T$

- Solving this system *once* provides the full specification of uncertain model ouputs

Laminar 2D Channel Flow with Uncertain Viscosity

- Incompressible flow
- Gaussian viscosity PDF
  - $\nu = \nu_0 + \nu_1 \xi$
- Streamwise velocity
  - $\mathrm{v} = \sum_{i=0}^{P} \mathrm{v}_i \Psi_i$
  - v_0: mean
  - v_i: i-th order mode
  - $\sigma^2 = \sum_{i=1}^{P} \mathrm{v}_i^2 \left\langle \Psi_i^2 \right\rangle$

v_0 $\quad$ v_1 $\quad$ v_2 $\quad$ v_3 $\quad$ σ

Non-intrusive Spectral Projection (NISP) PC UQ

- *Sampling*-based
- Relies on black-box utilization of the computational model
- Evaluate projection integrals *numerically*
- For any model output of interest $\phi(\boldsymbol{x}, t; \lambda)$:

$$\phi_k(\boldsymbol{x}, t) = \frac{1}{\langle \Psi_k^2 \rangle} \int \phi(\boldsymbol{x}, t; \lambda(\boldsymbol{\xi})) \, \Psi_k(\boldsymbol{\xi}) p_{\boldsymbol{\xi}}(\boldsymbol{\xi}) d\boldsymbol{\xi}, \quad k = 0, \dots, P$$

- Integrals can be evaluated using
    - A variety of (Quasi) Monte Carlo methods
    - Quadrature/Sparse-Quadrature methods

1D H_2-O_2 SCWO Flame NISP UQ/Chemkin-Premix
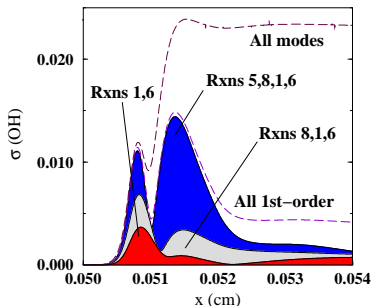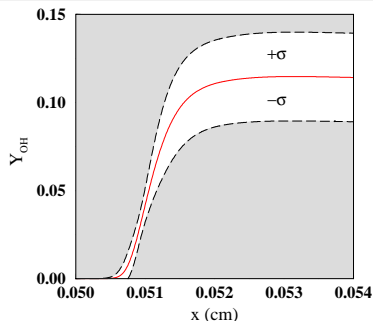
- Fast growth in OH uncertainty in the primary reaction zone
- Constant uncertainty and mean of OH in post-flame region
- Uncertainty in pre-exponential of Rxn.5 (H_2O_2+OH=H_2O+HO_2) has largest contribution to uncertainty in predicted OH

Chemical Model Reduction

- Chemical systems exhibit slow/fast behavior — stiffness
- They exhibit processes with fast time scales that are quickly exhausted, as the system decays onto low dimensional slow invariant manifolds (SIM), along which it evolves with the slow time scales
- The SIM is described by algebraic constraints

$$f_i(Y_1, \ldots, Y_N, T) = 0, \qquad i = 1, \ldots, M$$

  resulting from the equilibration of the fast processes

- These constraints can be used to determine a subset of the species in terms of the rest
- This equilibration is the basis for reduced-order behavior and model reduction

Introduction – Model Reduction Strategies

- Reviews (Griffiths, 1995; Okino & Mavrovouniotis, 1998)
- Lumping of variables
- Sensitivity and principal component analysis
- Reaction flux analysis (Frenklach 1986; Green 2003; Lu & Law, 2005)
- Quasi-steady state approx. (QSSA) (Hesstvedt 1978)
- Partial equilibrium approximation (Bulewicz 1956; Williams, 1985)
- Rate controlled equilibrium (RCCE) (Keck 1990; Law 1988)
- Computational singular perturbation (CSP) (Lam; Goussis 1988)
- Intrinsic low dimensional manifold (ILDM) (Maas & Pope 1992)
- Invariant constr. equil. preimage curve (ICE-PIC) (Pope 2006)

Computational Singular Perturbation (CSP)
Application to ODE system

- Jacobian eigenvalues provide first-order estimates of the time-scales of system dynamics: $\tau_i \sim 1/\lambda_i$
- Jacobian right/left eigenvectors provide first-order estimates of the CSP vectors/covectors that define decoupled fast/slow subspaces
- With chosen thresholds, have M "fast" modes
    - M algebraic constraints define a slow manifold
    - Fast processes constrain the system to the manifold
    - System evolves with slow processes along the manifold
- CSP *time-scale-aware* Importance indices provide means for elimination of "unimportant" network nodes and connections for a selected observable
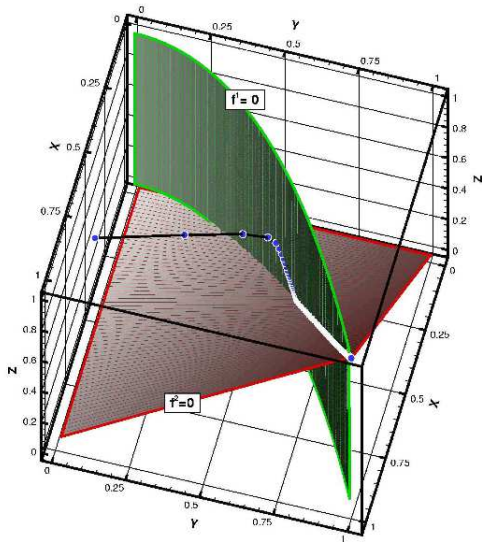
3D ODE System Example

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

$$\frac{d\mathbf{y}}{dt} = \mathbf{g} = \begin{bmatrix} -\frac{5\,y_1}{\varepsilon} - \frac{y_1 y_2}{\varepsilon} + y_2 y_3 + \frac{5\,y_2^2}{\varepsilon} + \frac{y_3}{\varepsilon} - y_1 \\[2ex] 10\frac{y_1}{\varepsilon} - \frac{y_1 y_2}{\varepsilon} - y_2 y_3 - 10\frac{y_2^2}{\varepsilon} + \frac{y_3}{\varepsilon} + y_1 \\[2ex] \frac{y_1 y_2}{\varepsilon} - y_2 y_3 - \frac{y_3}{\varepsilon} + y_1 \end{bmatrix}$$
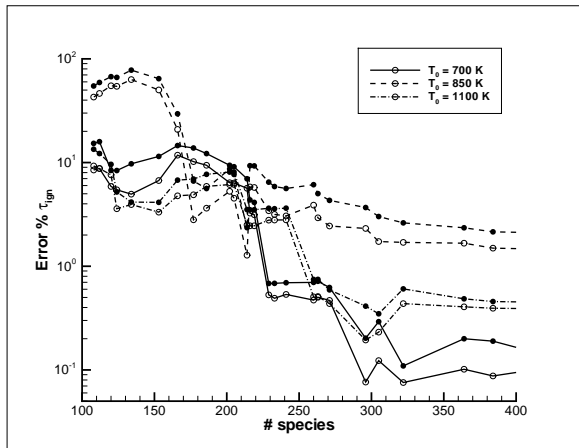
$\varepsilon << 1$: small parameter; controls the stiffness of the system

3D ODE System Dynamical Structure

- From any initial condition:
- System cascades through 2D, 1D manifolds to equilibrium

nHeptane Kinetic Model Simplification with CSP

- % Relative error in ignition time vs. simplified model sizes
- Control using error tolerances on CSP importance indices

Model Reduction under Uncertainty

Given uncertainty in the **detailed model** and its **parameters**:
It is of interest to

- Derive reduced models that are "good" over the range of uncertainty in the detailed model
- Constrain the error-level required from the reduced model per the given uncertainties in the detailed model

How do we

- Analyze system dynamics
  - Define manifolds, fast & slow subspaces
- Define measures of *importance* of reactions/species
- Define measures of *goodness* of a reduced model
- *Compare* uncertain models

Probabilistic Analysis of Uncertain ODE Systems

- Handle uncertainties using probability theory
- Every random instance of the uncertain inputs provides a "sample" ODE system
  - Uncertainties in fast subspace lead to uncertainty in manifold geometry
  - Uncertainties in slow subspace lead to uncertain slow time dynamics
- Probabilistic measures of importance
- Probabilistic comparison of models

- One can analyze/reduce each system realization
  - Statistics of $x(t; \lambda)$ trajectories
- This can be expensive $\quad \Rightarrow \quad$ explore alternate means

Intrusive Galerkin PC ODE System

$$\boxed{\frac{du}{dt} = f(u; \lambda)}$$

$$\lambda = \sum_{i=0}^{P} \lambda_i \Psi_i \qquad u(t) = \sum_{i=0}^{P} u_i(t) \Psi_i$$

$$\boxed{\frac{du_i}{dt} = \frac{\langle f(u; \lambda) \Psi_i \rangle}{\langle \Psi_i^2 \rangle} \qquad i = 0, \dots, P}$$

Say $f(u; \lambda) = \lambda u$, then

$$\frac{du_i}{dt} = \sum_{p=0}^{P} \sum_{q=0}^{P} \lambda_p u_q C_{pqi}, \quad i = 0, \cdots, P$$

where the tensor $C_{pqi} = \langle \Psi_p \Psi_q \Psi_i \rangle / \langle \Psi_i^2 \rangle$ is readily evaluated

Dynamical Analysis of the Galerkin PC ODE System

Key questions:

- How do the eigenvalues and eigenvectors of the Galerkin system relate to those of the sampled original system
- What can we learn about the sampled dynamics of the original system from analysis of the Galerkin system
  - fast/slow subspaces
  - slow manifolds
- Can CSP analysis of the Galerkin system be used for analysis and reduction of the original uncertain system

Stochastic system Jacobian: J

Reformulated Galerkin system Jacobian: $\mathcal{J}^P$

Key Results

1. The spectrum of $\mathcal{J}^P$ is contained in the convex hull of the essential range of the random matrix J.

$$\text{spect}(\mathcal{J}^P) \subset \text{conv}(\widetilde{W}(J))$$

2. As $P \to \infty$, the eigenvalues of $\mathcal{J}^P(t)$ converge weakly, *i.e.* in the sense of measures, toward $\bigcup_{\omega \in \Omega} \text{spect}(J(\omega))$.

3. $\mathcal{J}^P$ eigenvalues and eigenpolynomials can be used to construct polynomial approximations of the random eigenvalues and eigenvectors.

Sonday *et al.*, *SISC*, 2011; Berry *et al.*, in review

1D Linear Example
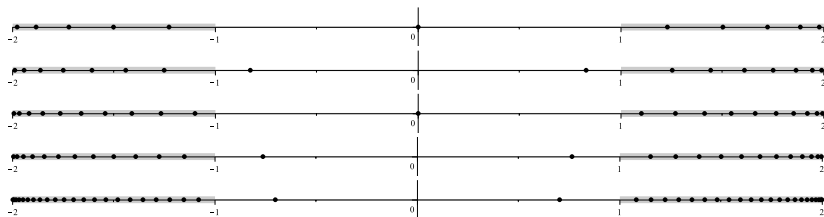
$$\dot{x}(\xi, t) = a(\xi)x(\xi, t); \quad \xi(\omega) \sim U[-1, 1];$$

$$J = a(\xi) \equiv \begin{cases} \xi + 1 & \text{for} \quad \xi \geq 0, \\ \xi - 1 & \text{for} \quad \xi < 0. \end{cases}$$

$$\widetilde{W}(J) = [-2, -1] \cup [1, 2]; \qquad \text{conv}(\widetilde{W}(J)) = [-2, 2].$$

LU PC: eigenvalues of $\mathcal{J}^P$ shown for $P = 10, 15, 20, 25, 45$

A 3D Non-Linear example

Makeev et al., JCP, 2002

$$\dot{u} = az - cu - 4duv \qquad\qquad \dot{v} = 2bz^2 - 4duv$$
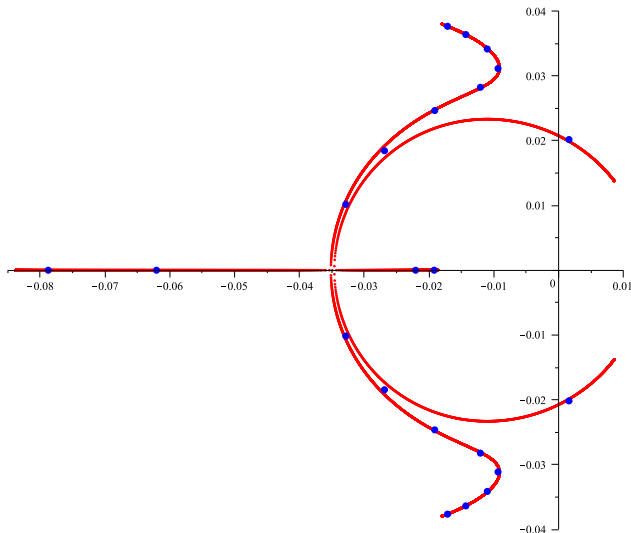$$\dot{w} = ez - fw \qquad\qquad z = 1 - u - v - w$$
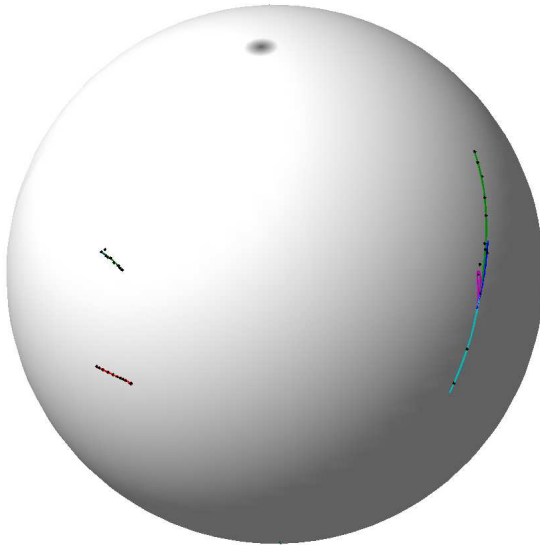
$$a = 1.6,\ b = 20.75 + .45\xi,\ c = 0.04$$
$$d = 1.0,\ e = 0.36,\ f = 0.016$$

$$u(0) = 0.1, v(0) = 0.2, w(0) = 0.7$$

3D Non-Linear Example; PC order 10

3D Non-Linear Example; PC order 10. Eigenvectors.

Closure

- Probabilistic uncertainty quantification (UQ) framework
  - Bayesian inference
  - Polynomial Chaos representation of random variables
    - Utility in forward/inverse UQ
- Chemical model reduction under uncertainty
  - Probabilistic framework
  - Eigenanalysis of the intrusive Galerkin PC system
    - Relationship to the uncertain system dynamics
- Work in needed on
  - Formulation of uncertain/stochastic manifolds
  - Stochastic importance criteria
  - Stochastic model comparison strategies
  - Uncertain model simplification algorithms
    - stochastic & kinetic dimensionality